

PAPER • OPEN ACCESS

Reconstructing complex system dynamics from time series: a method comparison

To cite this article: Forough Hassanibesheli *et al* 2020 *New J. Phys.* **22** 073053

View the [article online](#) for updates and enhancements.



PAPER

OPEN ACCESS

RECEIVED
31 March 2020REVISED
27 May 2020ACCEPTED FOR PUBLICATION
15 June 2020PUBLISHED
28 July 2020

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the
title of the work, journal
citation and DOI.



Reconstructing complex system dynamics from time series: a method comparison

Forough Hassanibesheli^{1,2,6} , Niklas Boers^{1,3,4} and Jürgen Kurths^{1,2,5}¹ Department of Physics, Humboldt-Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany² Research Domain IV—Complexity Science, Potsdam Institute for Climate Impact Research, Telegrafenberg A31, 14473 Potsdam, Germany³ Department of Mathematics and Computer Science, Free University Berlin, Germany⁴ Department of Mathematics and Global Systems Institute, University of Exeter, United Kingdom⁵ Saratov State University, 83, Astrakhanskaya Str., 410012 Saratov, Russia⁶ Author to whom any correspondence should be addressed.E-mail: hassanif@hu-berlin.de**Keywords:** complex systems, stochastic time series, Langevin equation, generalized Langevin equation, data-driven stochastic differential equation models

Abstract

Modeling complex systems with large numbers of degrees of freedom has become a grand challenge over the past decades. In many situations, only a few variables are actually observed in terms of measured time series, while the majority of variables—which potentially interact with the observed ones—remain hidden. A typical approach is then to focus on the comparably few observed, macroscopic variables, assuming that they determine the key dynamics of the system, while the remaining ones are represented by noise. This naturally leads to an approximate, inverse modeling of such systems in terms of stochastic differential equations (SDEs), with great potential for applications from biology to finance and Earth system dynamics. A well-known approach to retrieve such SDEs from small sets of observed time series is to reconstruct the drift and diffusion terms of a Langevin equation from the data-derived Kramers–Moyal (KM) coefficients. For systems where interactions between the observed and the unobserved variables are crucial, the Mori–Zwanzig formalism (MZ) allows to derive generalized Langevin equations that contain non-Markovian terms representing these interactions. In a similar spirit, the empirical model reduction (EMR) approach has more recently been introduced. In this work we attempt to reconstruct the dynamical equations of motion of both synthetic and real-world processes, by comparing these three approaches in terms of their capability to reconstruct the dynamics and statistics of the underlying systems. Through rigorous investigation of several synthetic and real-world systems, we confirm that the performance of the three methods strongly depends on the intrinsic dynamics of the system at hand. For instance, statistical properties of systems exhibiting weak history-dependence but strong state-dependence of the noise forcing, can be approximated better by the KM method than by the MZ and EMR approaches. In such situations, the KM method is of a considerable advantage since it can directly approximate the state-dependent noise. However, limitations of the KM approximation arise in cases where non-Markovian effects are crucial in the dynamics of the system. In these situations, our numerical results indicate that methods that take into account interactions between observed and unobserved variables in terms of non-Markovian closure terms (i.e., the MZ and EMR approaches), perform comparatively better.

1. Introduction

A complex system is a system composed of highly interconnected components in which the collective property of an underlying system cannot be described by dynamical behavior of the individual parts alone.

That is to say, the emergent behavior of a system is not evident in the components by themselves, but only in the system as a whole. Examples of this include lasers, fluids, financial markets, biological and social systems, but also Earth system and in particular climate dynamics variables [1–4]. Typically, complex systems are governed by nonlinear interactions and intricate fluctuations, and to retrieve the dynamics of a system from time series of only a few observed variables, it is required to characterize and assess interactions between deterministic tendencies and random fluctuations.

For complex systems, with large numbers of degrees of freedom interacting on various time scales, deriving explicit time-evolution equations remains both conceptually and computationally challenging. Many approaches from time series analysis and stochastic modeling have been proposed to model the behavior of complex systems based on observed time series by separating the system's behavior between observed macroscopic and hidden microscopic scales. Furthermore, it is typically assumed that these scales can be separated into slow and fast dynamics, respectively. Stochastic differential equations (SDEs) are then used as a result of the impossibility to determine all necessary processes and scales in comprehensive numerical models. In SDEs the macroscopic variables are modeled explicitly while the microscopic ones are represented by noise terms. The intuition underlying such approaches is to identify a small number of variables which are assumed to characterize the underlying dynamics, and solve the governing motion equation for the reduced-dimensional system in terms of SDEs. For instance master stability function has been extensively used to analyze the stability of the synchronous state in dynamical systems (both stochastic and deterministic) [5, 6].

During recent years, interest in modeling stochastic dynamics using Langevin equations (LEs) has elevated in many disciplines [7], such as investigating turbulent cascades [8], describing nano-friction fluctuations [9], or modeling molecular dynamics [10] and financial systems [11]. A commonly used, non-parametric method for deriving the deterministic drift and stochastic diffusion terms of an LE is based on the Kramers–Moyal (KM) coefficients, which correspond to the first and second order moments containing the spatial and temporal dependencies of the underlying dynamics [12, 13].

In the LE-approach, the dependencies between slow and fast variables are assumed to be negligible as a consequence of the separation of time scales. However, this assumption is not always well-grounded, for instance in Brownian motion [14], when the mass of particles is comparable with that of the surrounding particles, the random force does not obey white noise behavior anymore. Generalized Langevin equations (GLEs) have been proposed as a modification of LEs with no constraint of time scale separability [15]. A GLE equation can be retrieved from times series using projection operator techniques [15–17] and linear response theory [12]. In this study, we obtain the GLE model of a system within the projection operator framework. This technique, developed by Mori [15] and Zwanzig [18] (MZ), provides a mathematically consistent foundation for closed, reduced-dimension models. It is restrained, however, to Hamiltonian dynamical systems around thermodynamic equilibrium. This formalism applies a projection operator P that projects the whole system dynamics onto the subspace of slow variables, and projects out the remaining degree of freedom. As a consequence of this dimension reduction, non-Markovian terms arise explicitly in the resulting equation of motion, representing interactions between the slow (observed) and the fast (hidden) variables.

Another dimension-reduction approach, belonging to a class of multivariate stochastic models, is empirical model reduction (EMR) [19, 20]. This methodology, which can be understood as an extension of the classical linear inverse model (LIM) approach [21], has been extensively used in modeling observed climate variables [22] and in several cases can successfully reproduce their statistical characteristics. In constructing the EMR model, it is assumed that the time evolution of a system can be approximated by information embedded in the observed time series with only a coarse knowledge about a system's dynamics.

When fitting a prescribed drift term with free parameter to a real-world series, the residual will typically not be white neither in space nor in time; i.e., it will be state-dependent and exhibit serial correlations. To cope with this problem, the EMR approach iteratively models the time derivative of the residual as a linear function of the state of the observed variable and the residual itself, until the final residual is approximately independent (see methods section below). Substantively, this stochastic model constructs the evolution of macroscopic variables by introducing hidden components representing microscopic variables. In the parameterization procedure, hidden variables are approximated by the past observed variables, which constitutes a non-Markovian closure [20].

The main purpose of this work is to systematically investigate the statistical properties of different dynamical systems using the three aforementioned methods (KM, MZ, and EMR). In section 2, we shortly introduce the employed reduction methods. In section 3 we present a detailed comparison of the performance of these methods based on different kinds of synthetic and real-world time series, and discuss them briefly. Concluding remarks are provided in section 4.

2. Methods

2.1. Langevin equation

In the early twentieth century Paul Langevin proposed a quantitative description of random motion of colloidal particles suspended in a fluid, which has been known as a key problem of non-equilibrium statistical mechanics. This theory's applicability later has been extended to express the dynamical behavior of varieties of macroscopic systems without genuine particle ontologies. The original Langevin equation is a stochastic differential equation, which represents the time evolution of a subset of degree of freedoms containing both frictional and random forces that are associated with the fluctuation–dissipation theorem (FDT) [12].

Consider a system for which the evolution of the macroscopic states $x(t)$ obeys the following equation of motion:

$$\frac{dx(t)}{dt} = f(x, t) + g(x, t)\eta(t), \quad (1)$$

where $f(x, t)$ and $g(x, t)\eta(t)$ represent the deterministic force (e.g., friction and gravity) and stochastic forces (e.g., noise and chaotic particle interactions in many-body systems), respectively. Here, $\eta(t)$ is conventionally a stationary, δ -correlated Gaussian process with zero mean: $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t')\rangle = \delta(t - t')$. The presence of δ -correlated noise indicates that the Langevin process is a Markov process. For the class of stochastic processes following equation (1), the drift f and the diffusion g can be directly estimated from the measured data without any prior knowledge about the internal dynamics of a system using the Kramers–Moyal (KM) coefficients [23, 24]. The n th KM coefficient $D^n(x, t)$ can be understood in terms of the transition probability densities in the limit of $dt \rightarrow 0$. Formally, the KM coefficients are defined as:

$$D^n(x, t) = \frac{1}{n!} \lim_{dt \rightarrow 0} \frac{1}{dt} \langle (x(t + dt) - x(t))^n |_{x(t)=x} \rangle. \quad (2)$$

In this study, to calculate conditional moments from time-discrete data, we employed the Nadaraya–Watson estimator [25, 26]:

$$\frac{1}{dt} \langle (x_{(i+1)dt} - x_{idt})^j |_{x_i=x} \rangle = \frac{\sum_{i=1}^n K\left(\frac{(x_{idt}-x)}{h}\right) (x_{(i+1)dt} - x_{idt})^j}{\sum_{i=1}^n K\left(\frac{(x_{idt}-x)}{h}\right) dt}. \quad (3)$$

The kernel function K is here assumed to be Gaussian and accordingly the bandwidth h is determined using Silverman's rule of thumb: $h = 1.06\sigma N^{-\frac{1}{5}}$ where σ is the standard deviation of the time series under investigation.

The numerical discretization of the LE, in Itô's interpretation of stochastic integration, is as follows:

$$dx(t) = D^1(x, t)dt + \sqrt{D^2(x, t)}dW(t), \quad (4)$$

where dW denotes the increments of a Wiener process (known as a stochastic process with stationary independent normally distributed increments). The numerical integration of the equation (4) can be easily implemented using the Euler–Maruyama scheme with accuracy of order $\sqrt{\Delta t}$.

$$x(t + \Delta t) = x(t) + D^1(x, t)\Delta t + \sqrt{D^2(x, t)}\Delta t\eta(t). \quad (5)$$

Although the KM method has been applied successfully in many fields, the accuracy of estimation can be affected by finite-time effects since the estimation of drift and diffusion coefficients is sensitive to the sampling frequency of the underlying time series [27, 28]. Using the adjoint Fokker–Planck equation (AFPE) [29], we can derive D^1 and D^2 analytically for example for an Ornstein–Uhlenbeck (OU) process, and compare to the discrete-time estimates to judge on their accuracy:

$$\frac{\partial}{\partial t}P^+ = D^1 \frac{\partial}{\partial x}P^+ + D^2 \frac{\partial^2}{\partial x^2}P^+. \quad (6)$$

With a special initial condition $p^+(x, 0) = (x - X)$ we can solve the AFPE analytically. For instance, in the case of an OU process (containing 10^6 points with sampling frequency of 10^{-2}) subjected to additive and multiplicative noise, respectively, we compared the first and second order KM coefficients with those obtained from AFPE. As can be concluded from figure 1, the values of the KM coefficients estimated from discrete times series according to the equation (2) and the corresponding exact solutions are in a good agreement. One can also use the second-order expansion for moments [30] to check the robustness of estimated drift and diffusion coefficients for above systems.

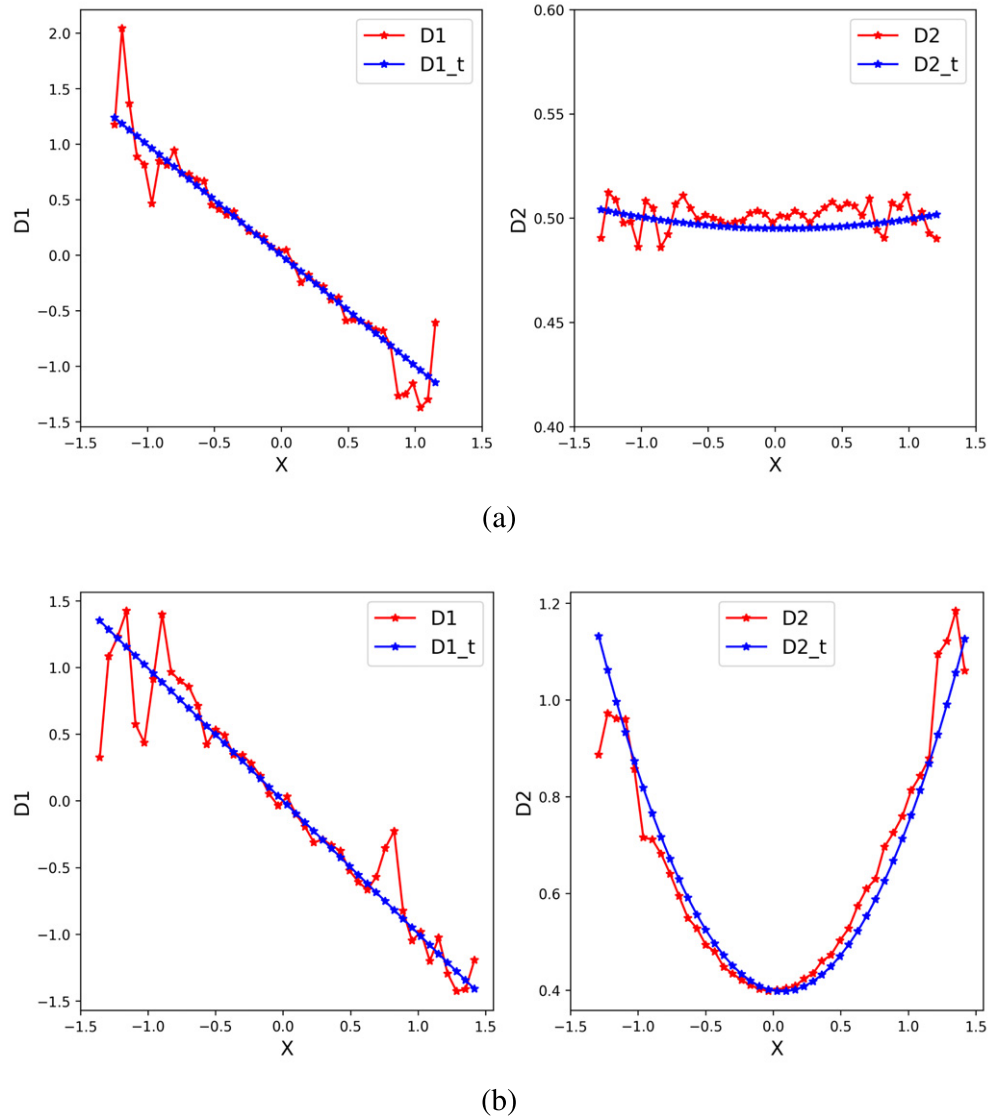


Figure 1. Comparison of the estimated drift ($D1$) and diffusion ($D2$) coefficients with corresponding exact solutions for an OU process containing 10^6 points with sampling frequency of 10^{-2} . (a) Estimated $D1$ and $D2$ (red) and the values obtained from the analytical solution via AFPE (blue) for an OU process with additive noise. (b) Same as (a) for an OU process subject to the multiplicative noise $1 + x^2$.

2.2. Generalized Langevin equation

The Langevin equation is the simplest approximation to describe the dynamics of non-equilibrium systems. As remarked earlier, the random force is assumed to behave like a Gaussian white noise, which does not lead to a good approximation if the time scale of the macroscopic variables is not much longer than that of the microscopic variables. This issue also requires a modification of the fluctuation dissipation theorem in which the correlation function of fluctuations is proportional to the memory of the frictional force. Therefore, generalized Langevin equations (GLE) have been proposed to account for long-range correlations and memory effects of complex systems that do not exhibit strong time scale separation. This generalized equation satisfies a certain fundamental consistency condition which relates the memory function to the auto-correlation of the stochastic force.

One way to derive a GLE is by means of the Mori–Zwanzig (MZ) formalism [17, 31], a projection-based dimension reduction method that redefines a set of ordinary differential equations into a reduced system with a time-independent Hamiltonian as long as the system is close to equilibrium. The MZ formalism assumes that a macroscopic system can be well described by projecting the full microscopic dynamics of a system onto the space of macroscopic variables. The resulting equation consists of three terms; the first, local term defines the self-interactions of the macroscopic variables, the second, non-local term describes memory dependencies of the macroscopic variables, and the last term represents the residual force

associated with the fast variables that is typically approximated by white noise:

$$\frac{dx(t)}{dt} = \Omega x(t) - \int_{t_0}^t K(t-t')x(t')dt' + R(t, t_0). \quad (7)$$

By multiplying both sides of equation (7) with $x(t_0)$ and taking the average [32–34], we arrive at the following expression:

$$\left\langle \frac{dx(t)}{dt}, x(t_0) \right\rangle = \langle \Omega x(t), x(t_0) \rangle - \int_{t_0}^t K(t-t') \langle x(t'), x(t_0) \rangle dt' + \langle R(t, t_0), x(t_0) \rangle \quad (8)$$

where $\langle R(t)x(t_0) \rangle = 0$, because the fast variable of the underlying system is uncorrelated with the slow variable. We can rewrite equation (8) in form of an Volterra integro-differential equation [35, 36] for the autocorrelation function $C(t) = \langle x(t), x(t_0) \rangle$:

$$\frac{dC(t)}{dt} = \Omega C(t) - \int_{t_0}^t K(t-t')C(t')dt'. \quad (9)$$

The memory kernel K can then be calculated by solving this equation with known dC/dt and C . Note that the discrete version of the integral $\int_{t_0}^t K(t-t')C(t')dt'$ is $h \sum_{j=0}^{n-1} K(t_{n-1}-t_j)C(t_j)$, where we applied the composite left rectangular formula, with $t_j = t_0 + hj$ using equal spacing h . Thus having obtained estimates for the deterministic non-Markovian and the memory term on the right-hand side of equation (7), we are able to determine the random force. The non-Markovian deterministic part on the right-hand side of the equation (7), corresponding to the drift, is nonlinear for systems not close to equilibrium and $\Omega x(t)$ is thus replaced with $F(x(t))$ with a nonlinear function $F(x(t))$. In such cases, we estimate this term using the first KM coefficient and estimate the memory kernel [34] via:

$$\frac{dC(t)}{dt} = \langle F(x(t)), x(t_0) \rangle - \int_{t_0}^t K(t-t')C(t')dt'. \quad (10)$$

Note that because we are considering discrete numerical time series, we are in practice using the following evolution equation:

$$x(n) = x(t_{n-1}) + h[F(x(t_{n-1})) - h \sum_{j=0}^{n-1} K(t_{n-1}-t_j)x(t_j) + R(t_{n-1})]. \quad (11)$$

2.3. Empirical model reduction

The primary goal of spatio-temporal modeling of complex systems is to characterize the properties of underlying dynamics with minimal assumptions about non-sufficient observations. Over the last few years, linear and nonlinear inverse stochastic modeling approaches have been intensively developed and applied to obtain reduced models that can explain the statistics of a full system [21, 37]. The linear inverse model (LIM) approach assumes that the relevant dynamics can be decomposed into linear deterministic and additive Gaussian random fluctuations terms. Although the linear simplification can be considered as a decent tool for describing dynamical systems close to equilibrium, not surprisingly, most of real-world systems exhibit nonlinear behavior. Hence, the empirical model reduction (EMR) [19] has been proposed as a nonlinear generalization of LIM approaches.

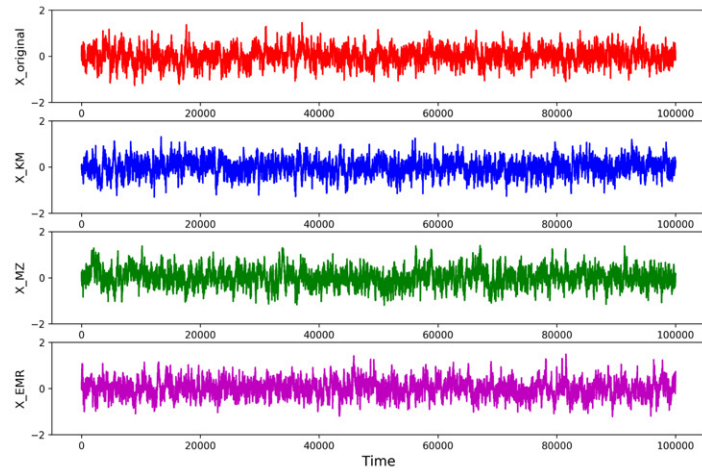
The general form of the EMR approach is as follows: first, model the increments $dx_k = x_{k+1} - x_k$ of an observed variable x as a nonlinear function of x plus a residual, typically via a least-squares minimization yielding the parameters of f :

$$dx_k = f(x_k)dt + r_k^0 dt \quad (12)$$

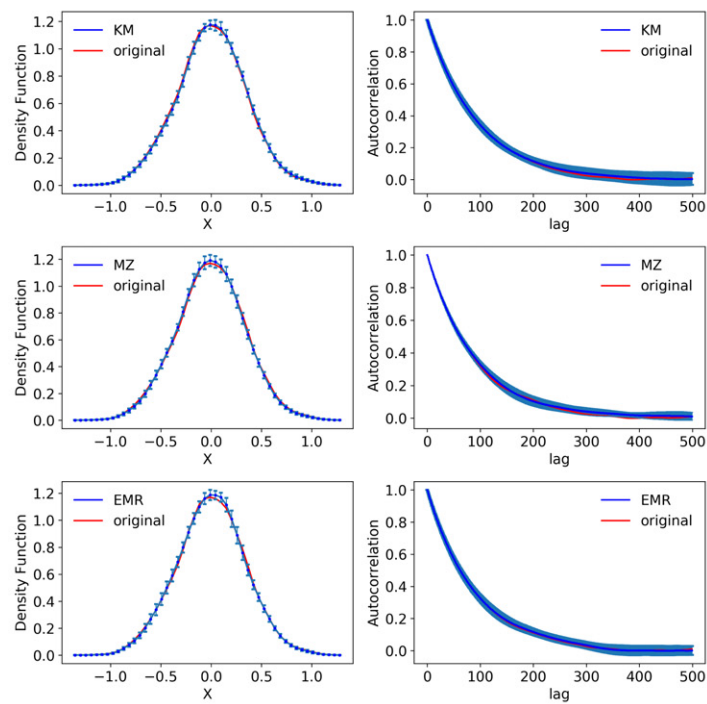
Second, model the residual as a linear function of x_k and r_k^0 and repeat this until the remaining residual has vanishing autocorrelation:

$$\begin{aligned} dr_k^0 &= b^1[x_k, r_k^0]dt + r_k^1 dt \\ \dots \dots \\ dr_k^n &= b^n[x_k, r_k^0, \dots, r_k^n]dt + r_k^{n+1} dt. \end{aligned}$$

Here, dr_k^n equals $r_{k+1}^n - r_k^n$ with time index k and n denoting the regression level. The presence of these hidden variables that explicitly depend on the past values of slow variables brings forth ‘memory’ effects. Incorporating these adaptive number of memory steps resembles the MZ formalism. Opposed to the MZ method, which provides solutions for the dynamics of a system considering the full memory via the kernel $K(t)$ function, the EMR offers approximate solutions based on the iterative procedure described above. We refer to Kondrashov *et al* [20] for further details and in particular explanations of the relationship between EMR and the MZ formalism.



(a)



(b)

Figure 2. A stochastic process with linear drift and additive noise. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom, respectively). (b) Summary statistics (PDFs in the left column and ACFs in the right column) of original and simulated time series derived from 1000 sample time series reconstructed by the three stochastic models (KM, MZ, and EMR), from top to bottom as indicated in the legend. The original system is an Ornstein–Uhlenbeck process (according to equation (13)) with $\Delta t = 0.01$ whose statistical features are shown in red color.

In equation (12), in fact, we substitute the deterministic term f by the first KM coefficient. This term can be a polynomial with different orders depending on the process dynamics. For instance, to reconstruct a stochastic process that takes place in a double-well potential $V(x)$, we first estimate the deterministic part via the first KM coefficient and then approximate the stochastic term using the recursive procedure.

3. Results and discussion

In the following, we present our results of KM, MZ, and EMR performance for reconstructing various dynamical systems with different levels of complexity. The performance of these three methods is examined in terms of statistical properties such as probability density functions (PDFs) and ACFs of large sets of time series simulated with the derived SDE models.

Table 1. Summary of mean squared error (MSE) between PDFs of different systems and corresponding averaged PDF of simulated time series over several realizations.

Time series	KM	MZ	EMR
OU additive noise	0.000 03	0.000 32	0.000 03
OU multiplicative noise	0.0013	0.0021	0.004
OU colored noise	0.002	0.000 01	0.000 01
S & P500	0.05	0.03	0.06
Niño-3	0.01	0.001	0.001
DW additive noise	0.0004	0.000 42	0.0021
DW multiplicative noise	0.0008	0.0009	0.001
[Ca ²⁺]	0.004	0.009	0.009

Table 2. The table illustrates MSE (as an error metric) between ACFs of original systems and corresponding average ACFs of simulated time series.

Time series	KM	MZ	EMR
OU additive noise	0.000 13	0.000 21	0.000 34
OU multiplicative noise	0.000 38	0.0003	0.000 27
OU colored noise	0.07	0.000 04	0.000 13
S & P500	0.002	0.001	0.0008
Niño-3	0.01	0.003	0.006
DW additive noise	0.000 15	0.000 02	0.000 31
DW multiplicative noise	0.0009	0.005	0.001
[Ca ²⁺]	0.026	0.004	0.004

3.1. Unimodal processes

3.1.1. Synthetic time series

We commence by considering the dynamics of the stationary Gaussian–Markov Ornstein–Uhlenbeck (OU) process:

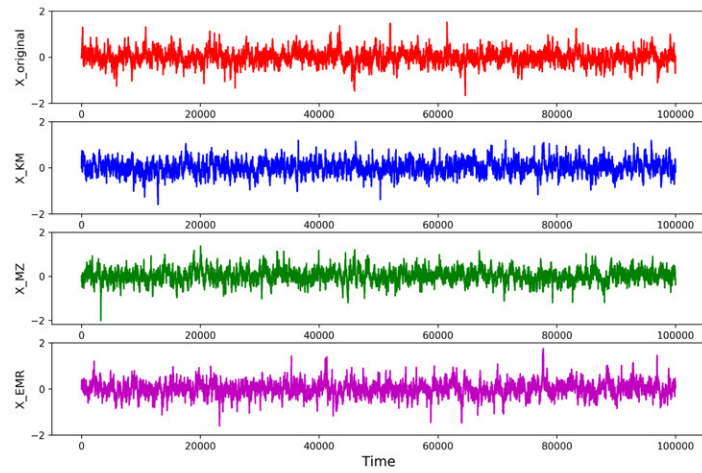
$$\frac{dx(t)}{dt} = \theta(\mu - x(t)) + \sigma\eta(t). \quad (13)$$

The OU process is known to be mean-reverting, i.e., the drift coefficient controls the forcing of the state variable $x(t)$ back to its mean μ . We generate 10^6 data points with $\theta = 1$, $\mu = 0$, $\sigma = 0.5$ and time sampling $\Delta t = 0.01$.

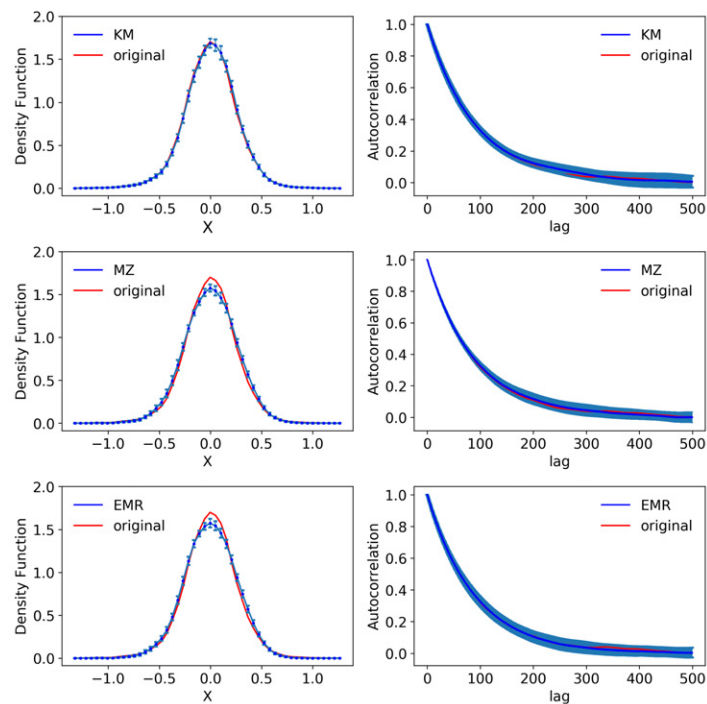
In order to numerically integrate equation (13), we apply the Euler–Maruyama method. We recall that for stochastic integration, there are two different approaches, Itô and Stratonovich. In the case of the OU process with an additive noise, these two interpretations are equivalent and in this study we mostly employ the Itô method.

In this study the deterministic terms of simulated time series constructed by the three aforementioned inverse modeling methods (KM, MZ, and EMR) are estimated directly from the first KM coefficient. Figure 2 displays the statistical properties of the time series generated using the original OU process as well as corresponding, randomly chosen time series obtained from the three different methods, in comparison with the original time series. The PDFs are obtained as averages over 1000 simulated time series and the error bars indicate the deviations $\pm\sigma$ around the averages for the 1000 realizations. We calculate the mean squared error (MSE) as an error metric to evaluate the performance of the KM, MZ, and EMR methods in reconstructing the statistical properties of underlying systems. The summary of MSE values can be found in tables 1 and 2. For the specific settings investigated in the case of equation (13), all three SDE models perfectly reproduce the statistic of the underlying dynamics. Here the auto-correlation exponentially decays and the memory coefficients of GLE are zero except at $\tau = 1$, indicative of a Markovian process.

Dynamical processes with multiplicative noise are ubiquitous in the natural sciences and beyond [38–40]. For a system subject to an additive noise the deterministic potential corresponds to stochastic steady-state potentials through $V_s = \frac{V_d}{\sigma}$, where σ is the intensity of fluctuations that are independent of the system state x . It implies that, in response to the stochastic force, the system exhibits stochastic fluctuations from its deterministic attractor. In the case of a multiplicative noise, a noise-induced drift becomes visible and consequently a new attractor basin is generated that does not exist in the absence of state-dependent fluctuations. In equation (13) we substitute the additive noise with a multiplicative noise (i.e., the diffusion term is multiplied by $1 + x^2$) whose intensity depends on the instantaneous value of the state



(a)



(b)

Figure 3. A stochastic process with linear drift and multiplicative noise $(1 + x^2)$. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Statistical properties of the observed and simulated Ornstein–Uhlenbeck subject to the multiplicative noise. The left column illustrates the PDFs for simulated results obtained as averages over 1000 realization with uncertainties in blue color. The right column displays the ACFs of ensembles of simulated time series constructed based on the three different stochastic models (KM, MZ, and EMR, from top to bottom).

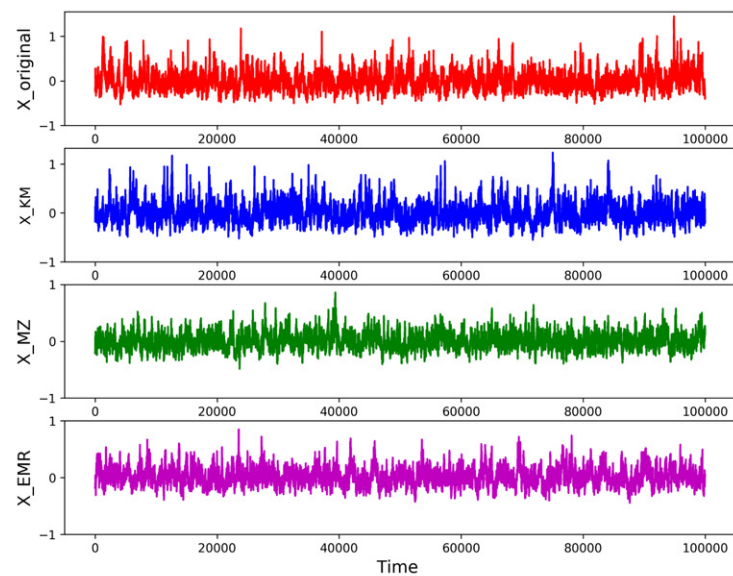
variable $x(t)$:

$$\frac{dx(t)}{dt} = \theta(\mu - x(t)) + (1 + x^2)\eta(t) \quad (14)$$

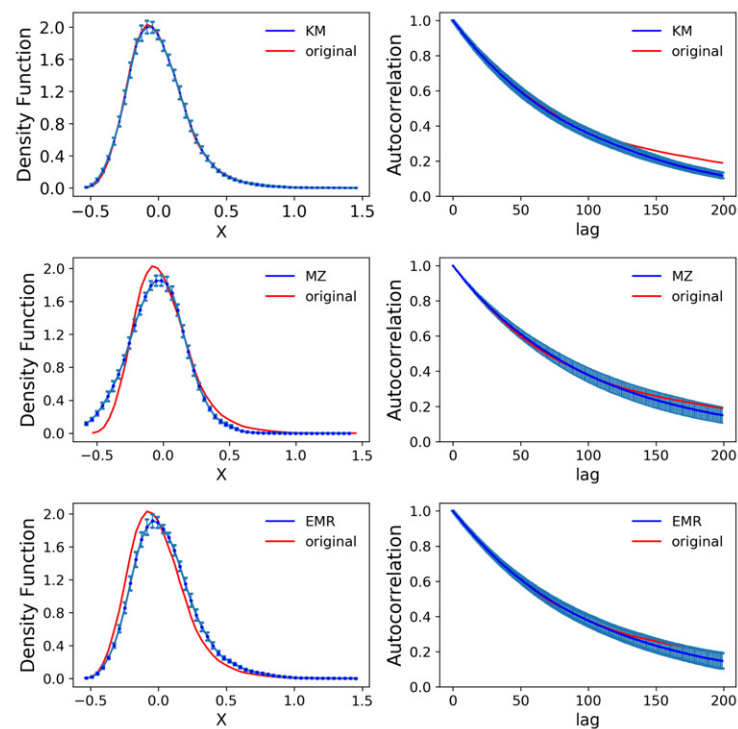
As it can be concluded from figure 3, all three methods work very well in reproducing the statistical features of the linear system also when exposed to symmetric multiplicative noise.

For more comparison, we analyze a system subject to asymmetric noise, which results in a skewed distribution. We thus multiply the noise term in equation (13) by $(1 + x)$ (instead of $(1 + x^2)$) and evaluate the statistical properties of the system. According to figure 4, the KM model exhibits better performance in comparison to the two other models, since it can directly estimate the (in this case asymmetric) state-dependent noise from the time series.

So far we considered additive and multiplicative noise, with η itself given by Gaussian white noise. But in many physical and biological systems, fluctuations exhibit some degree of correlation that cannot be



(a)



(b)

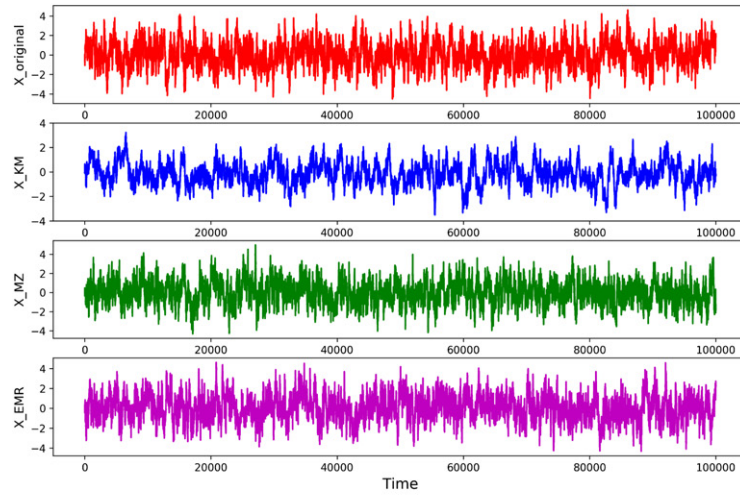
Figure 4. A stochastic process with linear drift and asymmetric multiplicative noise. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Statistical properties of the observed and simulated Ornstein–Uhlenbeck subject to the asymmetric noise $(1 + x)$. The left column illustrates the PDFs for simulated results obtained as averages over 1000 realization with uncertainties in blue color. The right column displays the ACFs of ensembles of simulated time series constructed based on the three different stochastic models (KM, MZ, and EMR, from top to bottom).

satisfied by white noise. Therefore we substitute the stochastic term of equation (13) with a first order autoregressive process (AR(1)) and investigate the performance of the three aforementioned methods to derive SDEs in the presence of colored noise.

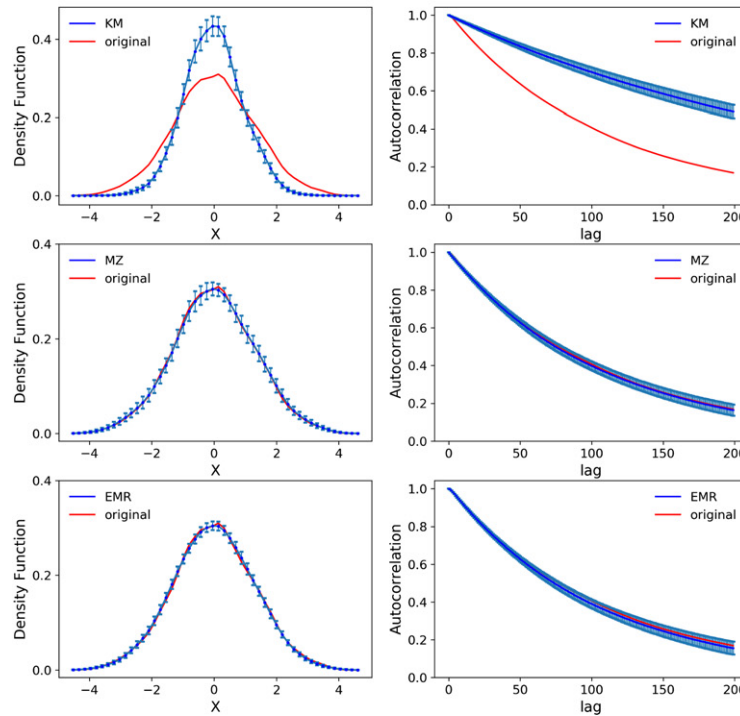
An AR(1) process is given by:

$$Y_{t+1} = \alpha Y_t + \eta(t), \quad (15)$$

where η is a Gaussian white noise process with zero mean and constant variance. According to figure 5(b), the statistical properties of simulated time series constructed via the KM method can not perfectly follow the original ones in this case. From tables 1 and 2 it can also be concluded that MZ and EMR perform



(a)



(b)

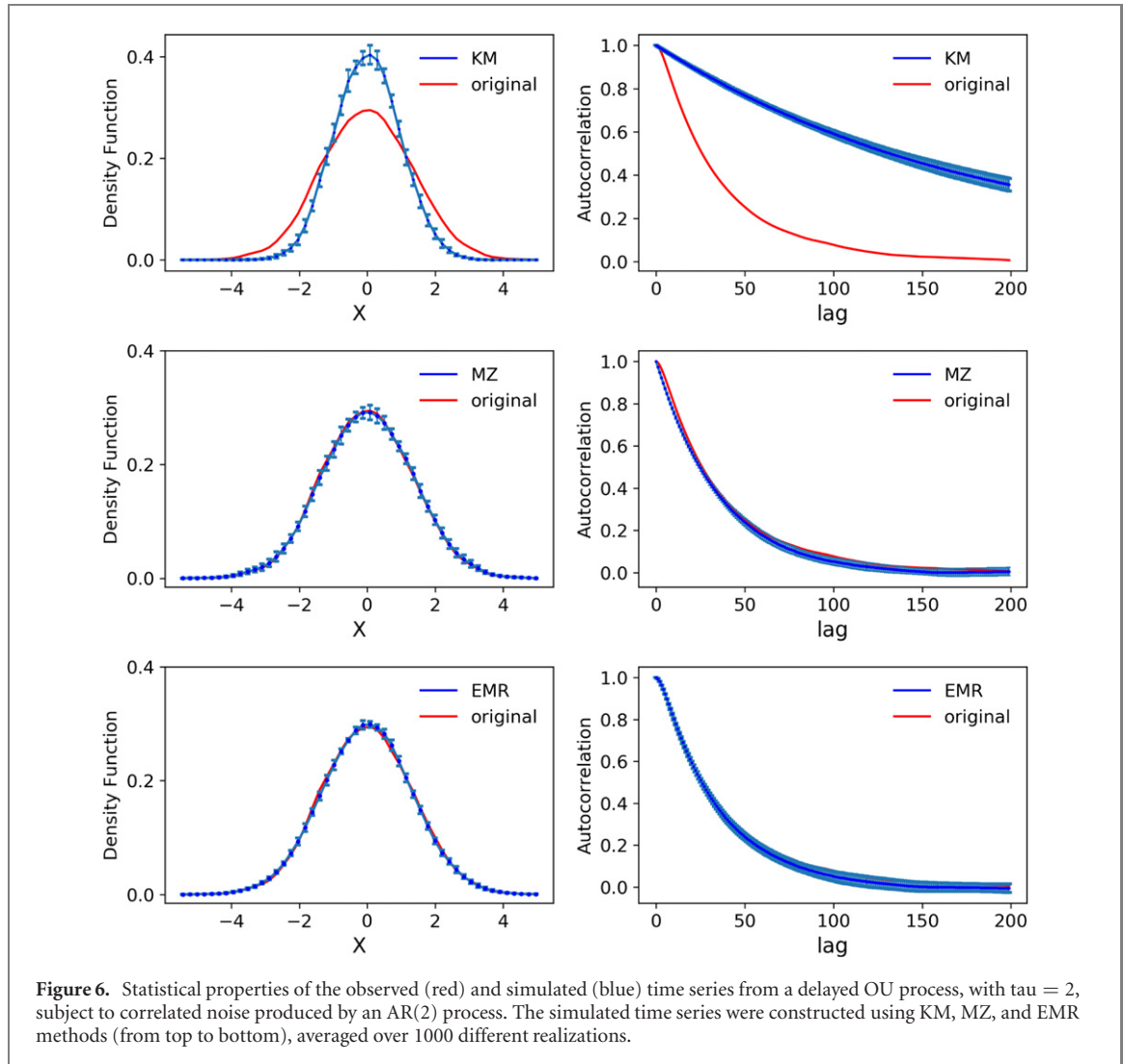
Figure 5. A stochastic process with linear drift and colored noise. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom). (b) Summary statistics (PDFs and ACFs) of original time series and simulated time series derived from 1000 sample time series reconstructed by the three stochastic models (KM, MZ, and EMR) from top to bottom as indicated in the legend. The original system is an Ornstein–Uhlenbeck process with colored noise constructed by a first-order autoregressive process (AR(1)) with $\alpha = 0.5$, whose statistical features are shown in red color.

better (with smaller MSEs). This is expected, since the presence of colored noise implies a deviation from the white-noise assumption of the LE and the KM method to derive it.

It is also of interest to evaluate the performance of these SDE models when the system exhibits some short-term memory in the deterministic part. For this purpose we consider a stochastic delay differential equation given by:

$$\frac{dx(t)}{dt} = \alpha x(t) + \beta x(t - \tau) + R(t). \quad (16)$$

The random force R is given by serially correlated noise produced by a second-order autoregressive (AR(2)) process in which the current value depends on the two previous values. As can be observed from figure 6, the EMR and MZ methods yield significantly better approximations of the PDF and ACF than the KM method for this system with short-term memory in the drift and colored noise. This provides an



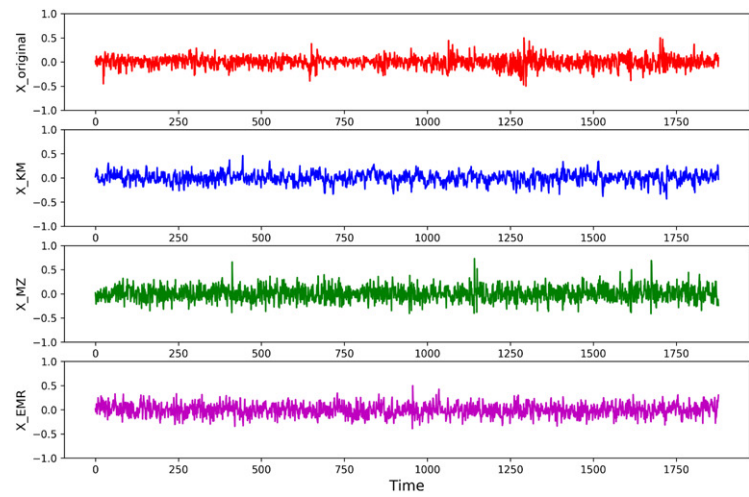
instructive example for situations where the MZ and EMR approaches clearly outperform the KM approach (as theoretically expected).

3.1.2. Real-world time series

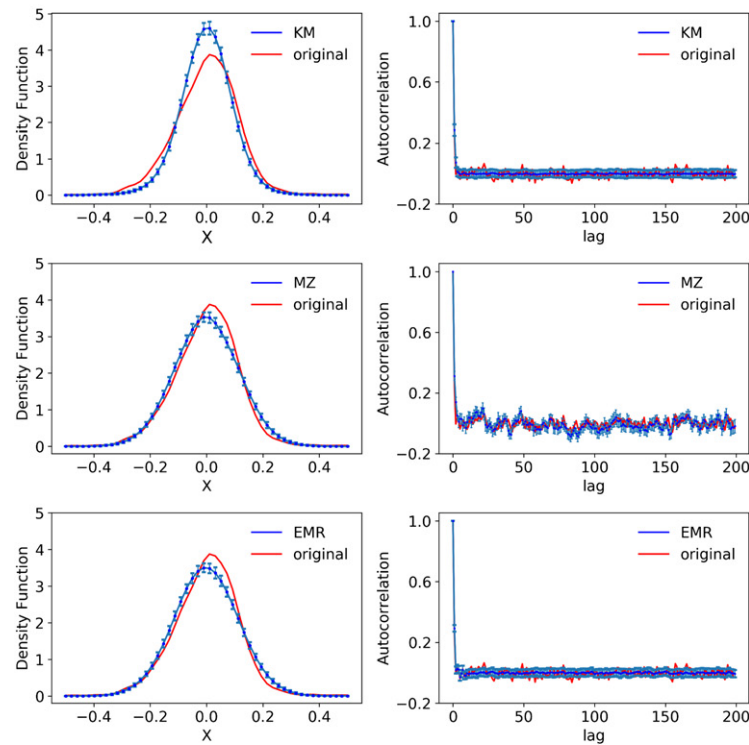
In recent years, quantifying stochastic dynamics of financial time series (e.g. stock prices and stock market indices) through stochastic differential equations in order to describe their evolution has attracted considerable attention. In this study we analyze the weekly S & P500 stock index for the time span of 35 years (1950–1985) and analyze its distinctive statistical properties. Here the stock return price $\ln \frac{p(t+\delta t)}{p(t)}$ represents the state variable $x(t)$. The statistical results of the simulated time series estimated by the three inverse modeling methods is shown in figure 7. According to the KM coefficients the deterministic part of the dynamics is described by a linear function of the state variable, while the stochastic term exhibits nonlinear behavior.

As mentioned above, we exert the first KM coefficient also to approximate the deterministic term of the GLE and EMR models. In this first real-world case, the approximation of the PDFs is not perfect anymore. The results reveal that the MZ formalism outperforms the two other methods in terms of the ACF, with very good agreement to the original return series (figure 7(b), second row, second column). It is worth mentioning that the statistical properties of financial time series exhibit a time-scale dependence and long-range correlations [41, 42]. Therefore we conclude that the memory effects of recent returns occurring in different time scales (from minutes to several days) can be modeled well by the MZ technique.

Another empirical time series that exhibits uni-modal variability is the Niño-3 index [43], which is one of several El Niño-Southern oscillation (ENSO) indicators in regard to tropical Pacific sea surface temperatures (SST). During the last decades, understanding the mechanisms underlying ENSO variability and prediction of future fluctuations has attracted substantial attention [44–47]. ENSO describes variations



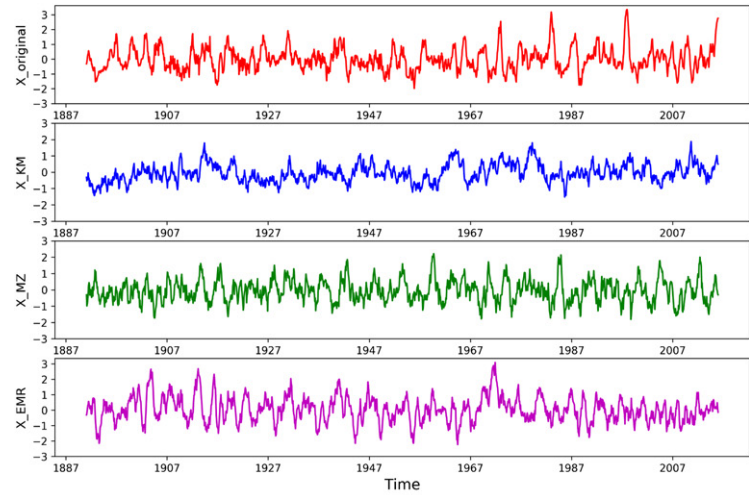
(a)



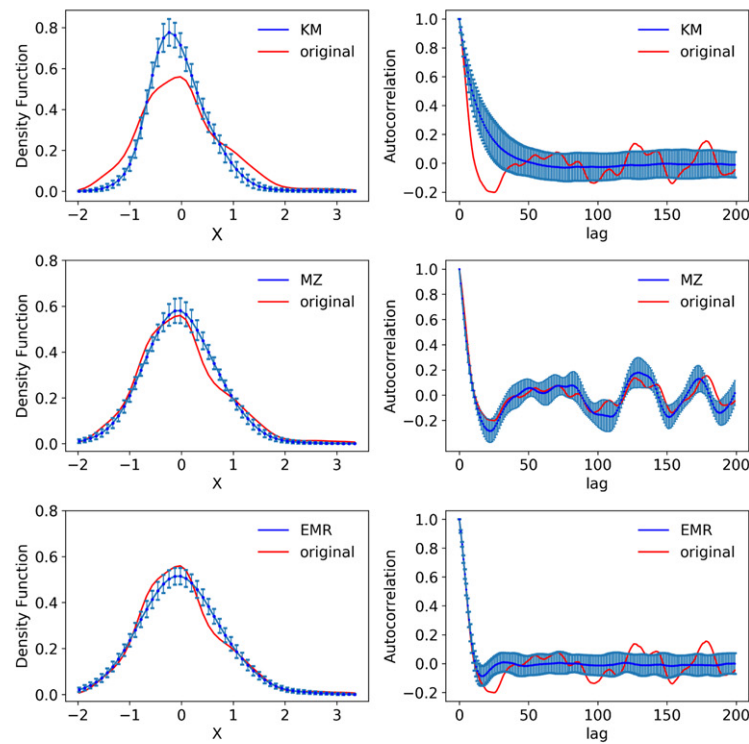
(b)

Figure 7. Summary statistics of S & P500 stock index. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) The average PDFs (left column) and ACFs (right column) of simulated time series over 1000 realizations (blue) created by KM, MZ, and EMR models from left to right, respectively. The statistics plotted in red correspond to the original weekly S & P500 stocks index for the time span of 35 years (1950–1985).

in temperature and pressure in the eastern Pacific ocean and has significant impacts on global climate variability. We reconstruct the Niño-3 monthly sea surface temperature (SST) indices averages across (5N–5S, 150–90W) from 1891 to 2015 using the KM, MZ, and EMR inverse modeling approaches. The non-Gaussian behavior of Niño-3 indicates a nonlinear process, quantified by the positive skewness of the SST distribution, which may reside in the interaction of oceanic variables of interest and the atmospheric fast forcing [48]. This kind of interactions brings about memory effects into the system dynamics that should not be expected to be fully captured by models established based on Markovian assumptions. Figure 8(b) displays the PDFs and ACFs of the resulting simulated time series. Although all three inverse methods produce similarly skewed and heavy-tailed distributions, comparing the ACFs, MZ achieves a higher accuracy approximation than KM and EMR. We repeated this procedure for the Niño-4 index, and also in that case found that the MZ approach outperforms the other two methods. According to these



(a)

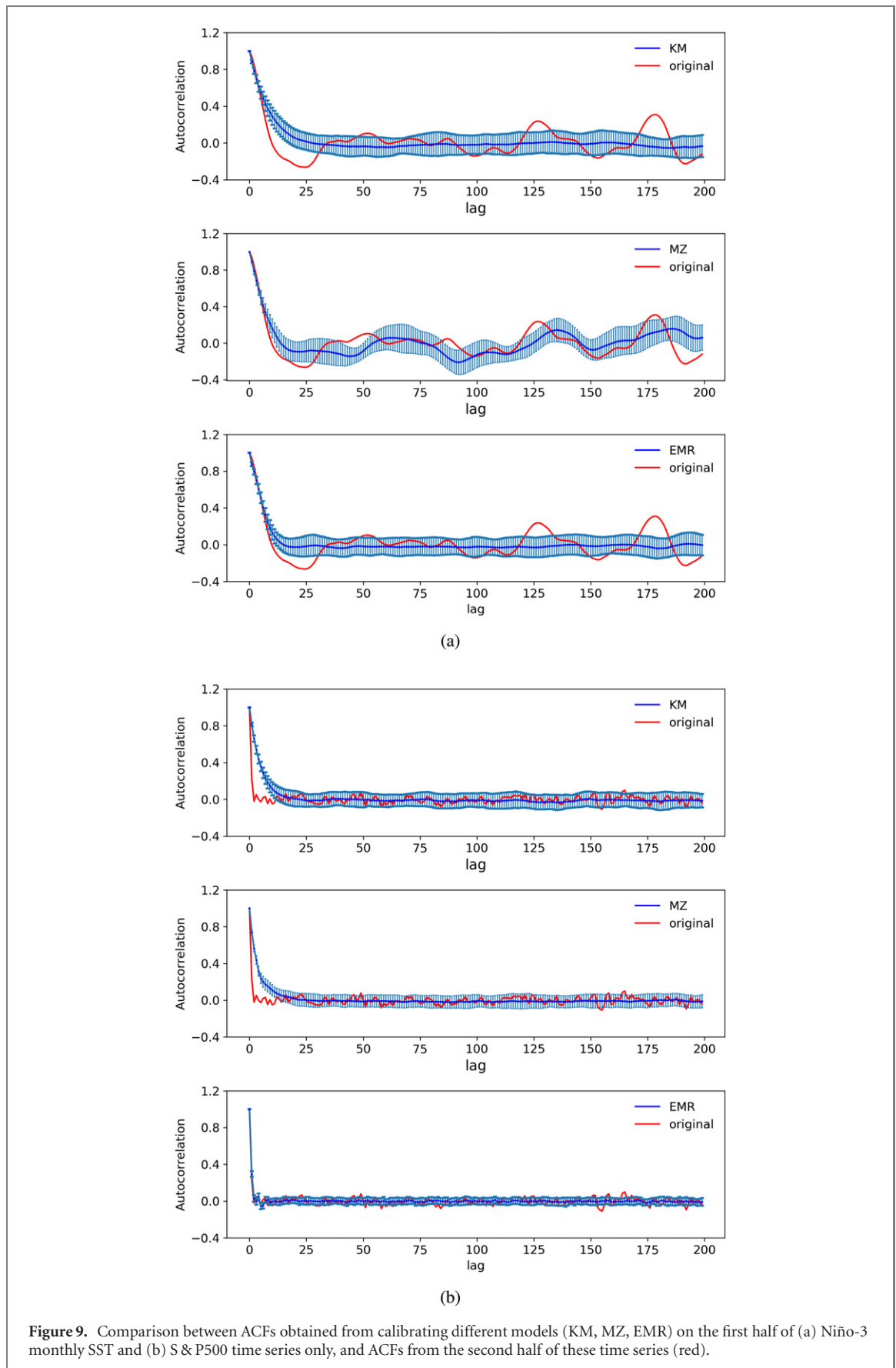


(b)

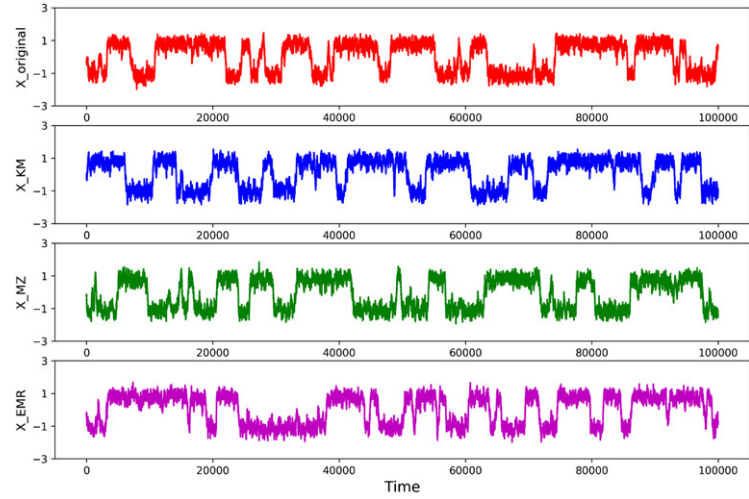
Figure 8. Summary statistics of Niño-3 monthly sea surface temperature (SST) indices. (a) Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Comparison of model performance of KM, MZ, and EMR in reconstructing dynamics underlying ENSO variability. The figure depicts the statistical properties of the observed (red) and simulated (blue) monthly Niño-3 SST indices from 1891 to 2015.

results, the most significant characteristic of these real-world processes leading to ENSO variability is the presence of serial correlations connected to internal interactions between observed (slow) and unobserved (fast) variables which can be captured best by the MZ approach.

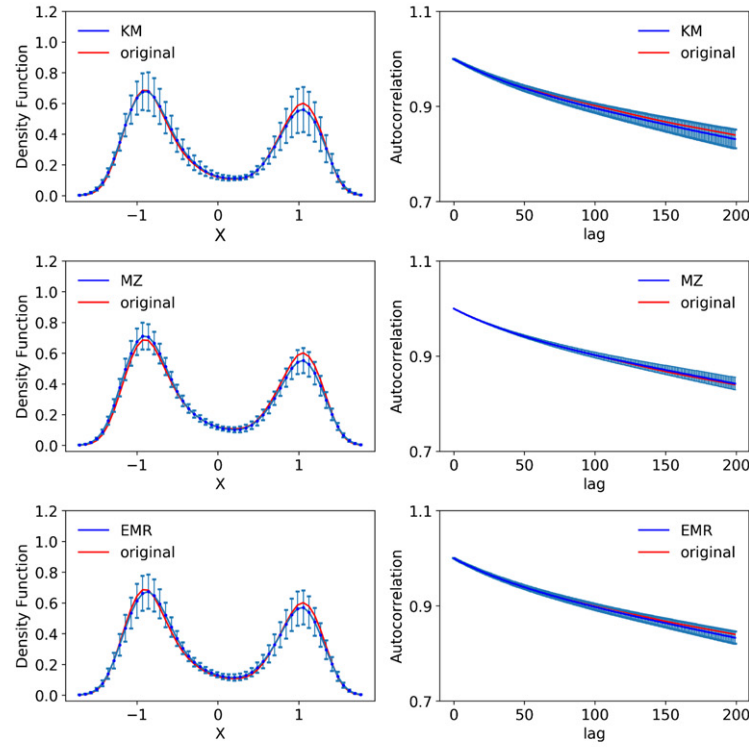
In order to investigate whether these results are prone to overfitting, we first calibrated the three SDE models on the first half of the time series and validated the parameters by comparing the resulting time series statistics with the ones of the second half of the time series for both systems (Niño-3 and S & P500, respectively). The results are presented in figures 9(a) and (b). It can be inferred that the longer-term variations in the ACF are still captured to some degree by the MZ method. For the ENSO case, the slow variations of the ACF correspond to low-frequency variability present in this time series, which is not noise but an important part of (internal) climate variability. The fact that the reproduction of these slow



variability modes is less accurate when only calibrating on the first half of the time series is due to the fact that capturing these slow variability modes is harder when considering only a shorter part of the time series. For the S & P500 time series however, the comparably faster fluctuations of the ACF most likely correspond



(a)



(b)

Figure 10. A stochastic process with nonlinear drift and additive noise. (a) Original (red) and simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Comparison of the numerical simulation of a process in a double-well potential subject to an additive Gaussian noise with intensity $\sigma = 0.5$, 10^6 data points and a sampling interval of 0.01. The plot displays PDFs and ACFs of the original (red) and simulated (blue) time series obtained from three different stochastic models (KM, MZ, and EMR as indicated in the legend), averaged over 1000 realizations.

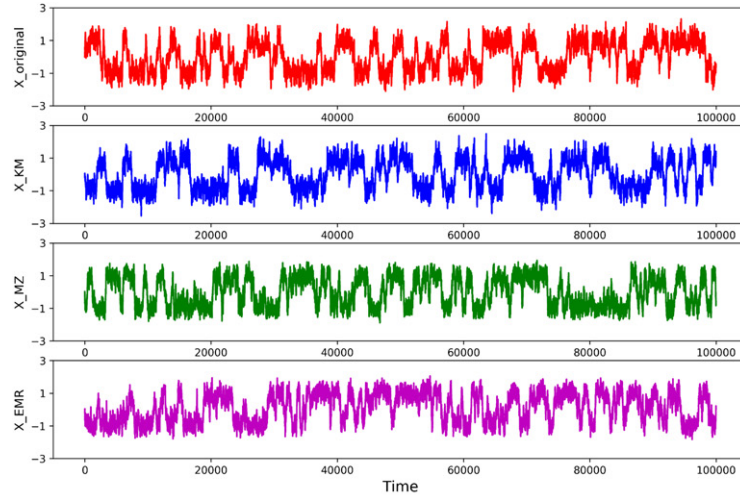
to noise. The very good approximation of the ACF when applying the MZ formalism to the full S & P500 time series could hence be the result of overfitting when deriving the full kernel K .

3.2. Bimodal processes

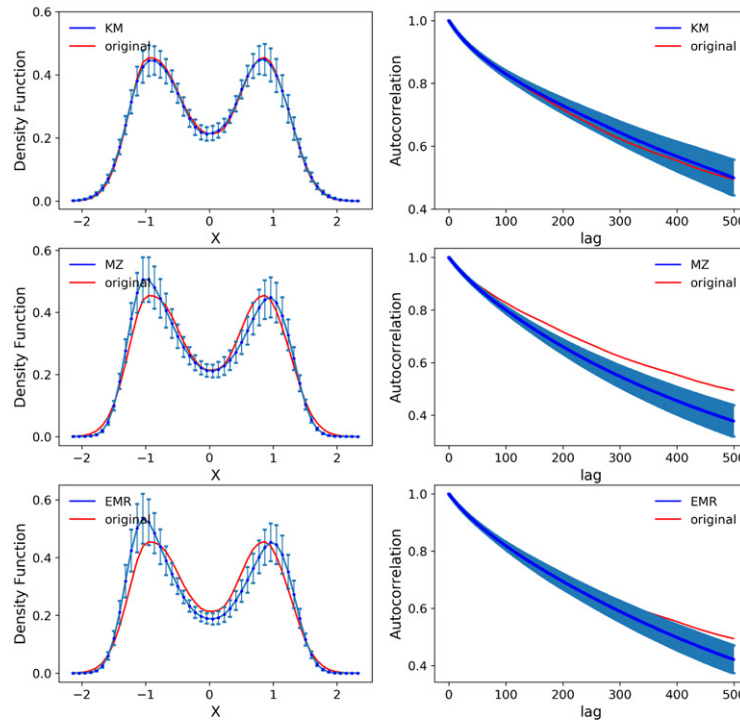
3.2.1. Synthetic time series

Up until now we studied processes with a unique mode; however, there are various stochastic dynamics in natural systems that do not exhibit uni-modality. We start with generating a synthetic process in which a particle moves in a double-well (DW) potential at $\pm\sqrt{\theta}$, driven by additive Gaussian noise

$$\frac{dX(t)}{dt} = \theta(x(t) - x^3(t)) + \sigma\eta(t). \quad (17)$$



(a)



(b)

Figure 11. A stochastic process with nonlinear drift and multiplicative noise. (a) Original (red) and simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Statistical properties of the observed (red) and simulated (blue) time series related to a particle motion in a double-well potential, subject to multiplicative noise with diffusion $\sqrt{1+x^2}$. The simulated time series were constructed using KM, MZ, and EMR methods (from top to bottom), averaged over 1000 different realizations.

Here we construct the stationary time series with $\theta = 1$, $\sigma = 0.5$ and $\Delta t = 0.01$. For all three inverse models the specific values of the drift terms are again derived from the first-order KM coefficients from the original time series. The non-Markovian term in the GLE derived via the MZ formalism, which accounts for interactions between the slow and fast variables, was estimated from the information encoded in the ACF [32, 34] (equation (7)). Considering the Markov property of the underlying system (17), the kernel function K of the MZ approach is zero (as theoretically expected). The left column in figure 10 illustrates the average PDFs of model-simulated time series in comparison with original time series, for 1000 realizations, and the right column displays the associated ACFs. It can be seen that these three stochastic models can mimic the key features of the original time series very well.

In contrast, in the presence of a multiplicative noise, estimating the underlying dynamic is not trivial anymore. We substitute the additive noise in the previous example with $\sqrt{1+x^2}$. As it can be observed in

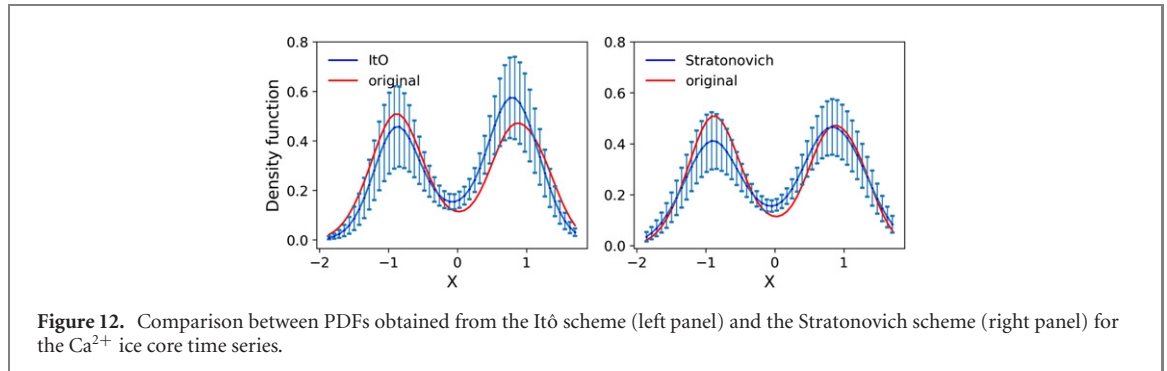


figure 11), the average ACFs deviate from the original ones for the MZ and EMR methods, while the KM method still achieves an excellent approximation of the PDF and ACF.

3.2.2. Real-world time series

Identification and understanding the triggering mechanism of the Dansgaard–Oeschger events, abrupt temperature increases during the last glacial period (roughly from 115.000 years to 12.000 years BP) has attracted considerable attention in the past years [49–53]. These rapid and strong warming events (up to 16K within few decades [54]), followed by slow-paced cooling phases are inferred from the content of stable isotope composition of water $\delta^{18}\text{O}$ in different Greenland ice cores. These proxy records allow to empirically reconstruct the temperature profile over the last 120.000 years with high temporal resolution. In this study we investigate the high-resolution (20 years-average) Ca^{2+} (interpreted as a proxy for atmospheric circulation patterns), collected from the NGRIP ice core on the GICC05 time scale [55]. Because of the substantially better signal-to-noise ratio, we focus here on the Ca^{2+} time series between 60ka and 30ka b2k. We apply the KM, MZ, and EMR approaches to reproduce dynamical and statistical properties of the underlying dynamics of $[\text{Ca}^{2+}]$.

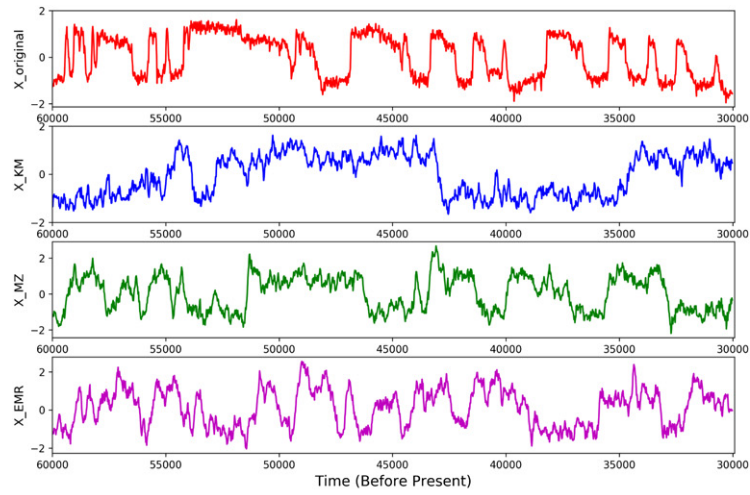
Since the integration of the diffusion term of an SDE is not uniquely defined, two different frameworks for stochastic calculus have been proposed, namely the Itô and the Stratonovich scheme (see appendix A for further details). Up to now, we have interpreted the SDEs in the Itô sense, for which a consistent discretization scheme is the Euler–Maruyama method with straightforward white-noise increments.

In contrast, in the Stratonovich calculus white noise is approximated by continuously fluctuating noise with finite memory, which may be more suitable for approximating real-world time series. Indeed, it has been argued that the Stratonovich scheme is more appropriate for continuous physical systems, such as ocean and atmospheric circulation systems [56]. A well-defined SDE discretization scheme that is consistent with the Stratonovich interpretation is the Heun method, which has been shown to perform with higher precision for a dynamical system model of paleoclimate variability [57].

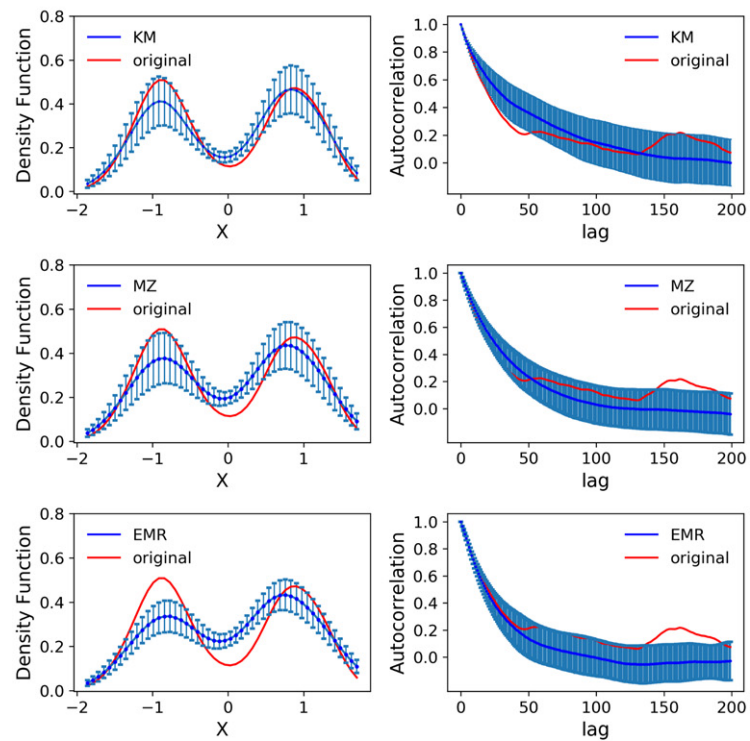
We compared the performances of both stochastic calculi and corresponding discretization schemes (Euler–Maruyama and Heun, respectively) for the Ca^{2+} ice core time series. According to figure 12, the Stratonovich calculus performs better than the Itô calculus in terms of approximating the PDF. This likely corresponds to the intrinsic properties of the Stratonovich interpretation in which fast quantities with finite correlation time, as observed in continuous physical systems such as climate variables, can be approximated as white noise. The better performance obtained using the Stratonovich calculus is in agreement with observations made in earlier studies [56, 57].

Figure 13 shows the statistical properties (PDFs and ACFs) of the observed and simulated $[\text{Ca}^{2+}]$ time series in the interval 60.000 years to 30.000 years before the year 2000 AD. Because of the large amplitude of the $[\text{Ca}^{2+}]$ concentration variations, the calculations were conducted in natural logarithmic scale [50]. The drift term, again directly derived from the KM coefficient for all three inverse methods, exhibits a nonlinear behavior with two stable equilibrium states, which reflects the switching mechanism between cold stadials and warmer interstadials.

The results presented in figure 13 show that the modeled time series for all three methods could accurately reproduce the bimodality of the observed PDFs, although it should be noted that the LE derived with the KM approach approximates the observed PDF better than the GLE and EMR approaches: the depth of the two potential wells in MZ and EMR are shallower than the observed PDF suggests. It is clear from the results illustrated in the right column of the figure 10(b) that the MZ



(a)



(b)

Figure 13. Comparison between observed and simulated $[\text{Ca}^{2+}]$ time series from Greenland ice cores. (a) Original (red) and simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively). (b) Statistical properties of the observed and simulated $[\text{Ca}^{2+}]$ time series in the interval between 60ka and 30ka b2k. The PDF of the original data is shown in red color while simulated time series obtained from KM, MZ, and EMR models (from top to bottom) can be found in blue color. The PDFs for the three stochastic models are averaged over 1000 realization and therefore considerably smoothed.

and EMR methods have better performances to construct the underlying autocorrelation structure of the $[\text{Ca}^{2+}]$.

4. Conclusions

In the last decades, there have been considerable methodological advances in understanding and quantitatively modeling the behavior of complex systems. Among others, stochastic differential equations (SDEs) are a promising approach for this challenge in situations where only a few variables are actually measured. For some typical example settings, we have studied the performance of three methods to derive SDEs—the Kramer–Moyal approach to derive Langevin equations (LE) with potentially multiplicative

noise, the Mori–Zwanzig approach to derive generalized Langevin equation (GLE) including a non-Markovian term, and the empirical model reduction (EMR) approach to derive GLEs—for various synthetic and real-world time series. Since the microscopic dynamics of complex systems are often not accessible, these methods regard the unobserved, fast variables as random fluctuations and investigate the properties of underlying systems from the evolution of only a few macroscopic, comparably slow observables. In this study corresponding numerical simulations of all three inverse methods have been examined in terms of PDFs and ACFs of the simulated time series, as metrics for assessing the models' performance. We generally observed a nearly optimal performance of all three approaches for unimodal Markovian systems. For unimodal, non-Markovian systems, the MZ and EMR strongly outperformed the KM approach as theoretically expected. It is thus not possible to draw a firm conclusion about which SDE deviation method is in general superior in comparison to the others. According to our results, the performance of SDE models strongly relies on the effects of memory on the underlying dynamics. We could show that LEs (as derived by the KM approach) obtain better results for systems with weak memory contribution but asymmetric multiplicative noise, since it can directly estimate the state-dependent noise with higher precision. On the other hand, for systems with memory effects and colored-noise forcing, it is essential to take into account the non-Markovian closure terms. Hence, the MZ and EMR approaches can be considered more reliable in reconstructing the dynamics of systems exhibiting strong memory effects. In these two methods, the interactions between observed and unobserved variables are taken into consideration in terms of memory effects. That is, the EMR approach incorporates adaptive numbers of memory steps, whilst the MZ method considers the full memory of a system via the kernel K . With regards to prospective investigations, we suggest that the adaptability of the three investigated SDE models in reproducing dynamical characteristics of multi-variate time series should be further investigated. For example, Boers *et al* [50] introduced a two-dimensional stochastic delay differential equation model as an approximation of the GLE to analyze the temperature and dust proxy records from Greenland ice cores. Their results revealed the role of the coupling strength between the two variables, where the model has problems in reproducing the bimodality in absence of coupling terms, and also shows that the explicit consideration of memory terms helps to reproduce the asymmetry observed in those temperature and dust proxy records. Furthermore, it should be of interest and high practical relevance for real-world systems to consider an adaptive stochastic differential equation framework that can reconstruct the underlying dynamics of complex systems driven by Lévy noise.

Acknowledgments

This work has been financially supported by the German Academic Exchange Service (DAAD) under project no. 57299294. NB acknowledges funding by the Volkswagen foundation. This is TiPES contribution #25; the TiPES (Tipping Points in the Earth System) project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 820970. JK was supported by the project RF Government Grant 075-15-2019-1885.

Appendix A. Stochastic calculus

Equation (1) with a given initial condition $x(t_0) = x_0$ has a unique solution which satisfies the following integral form:

$$x(t) = x(t_0) + \int_{t_0}^t f(x(t'))dt' + \int_{t_0}^t g(x(t'))\eta(t')dt'. \quad (\text{A1})$$

A Wiener process is non-smooth and nowhere differentiable, hence translating the second integral in the equation (A1) using a conventional Riemann sum is not uniquely defined. To interpret the noise term, two different formulations of stochastic calculus have been introduced for computing the solutions of SDEs; known as Itô and Stratonovich calculi, respectively [58, 59].

In the Itô prescription, the evaluations of the function $g(x(t))$ are uncorrelated with the (infinitesimal) increments of the Wiener process $dW(t)$. In fact, the Itô integral is defined as the limit of a left Riemann sum (where the function $g(x(t))$ is evaluated at the left of the interval $[t, t + \Delta t]$) with an Itô correction. The resulting stochastic integration in the Itô scheme is given:

$$\int_{t_0}^t W(s)dW(s)ds = \frac{1}{2}[(W^2(t) - W^2(t_0)) - (t - t_0)]. \quad (\text{A2})$$

Although the Itô convention does not preserve the chain rule of calculus, employing Itô's Lemma maintains the martingale property. Owing to this property, the Itô calculus is used extensively in finance [60].

The most common alternative to the Itô integral that does satisfy the chain rule of classical calculus is the Stratonovich scheme. From that point of view, a function can be evaluated at the midpoint of the time interval $[t, t + \Delta t]$. Because the midpoint selection rule is associated with the finite noise autocorrelation [61], the martingale property does not hold. In contrast to the Itô, the Stratonovich calculus approximates the Wiener process as the limit of a correlated process when the correlation time approaches zero:

$$\int_{t_0}^t W(s) \circ dW(s) = \frac{1}{2}(W^2(t) - W^2(t_0)). \quad (\text{A3})$$

This approximation leads to difficulties e.g. for the calculation of expectation values, since stochastic variables and noise are not independent: $\langle x(t)\eta(t) \rangle \neq 0$. It should be underlined that the Itô and Stratonovich calculi have the same solution if their drift terms fulfill the following relationship, which is called Itô–Stratonovich drift correction:

$$f_S(x(t)) = f_I(x(t)) - \frac{1}{2} (g(x(t)) \frac{\partial g(x(t))}{\partial x}), \quad (\text{A4})$$

where f_I denotes the drift of the Itô calculus and f_S the drift of the Stratonovich calculus.

Even though both interpretations are mathematically consistent, one obvious question that may arise is which interpretation is the right one for describing or approximating a particular set of physical processes. In order to answer this, we have to look at the origin of the noise in the stochastic system. It has been shown that, if the relaxation time of a system is large enough in comparison with the noise correlation time, then the Itô interpretation is appropriate. On the other hand, if the noise is colored, i.e., it has finite correlation time, the limiting SDE must be treated in the Stratonovich framework [61–63].

As noted above, different stochastic calculi (Itô or Stratonovich) are associated with different kinds of discretization for numerical integration [64]. It must hence be stated upfront which stochastic calculus is going to be considered. The simplest and most widely used discretization scheme to numerically integrate SDEs is the Euler–Maruyama method, which converges to the Itô interpretation:

$$x(t + \Delta t) = x(t) + f(x(t))\Delta t + g(x(t))\eta(t)\sqrt{\Delta t}. \quad (\text{A5})$$

Another numerical method that we used in this paper (for approximating the Ca^{2+} ice core time series), called Heun method [65], leads to the Stratonovich scheme. This method is an example of a predictor-corrector method in which the predictor is calculated by a simple Euler type integration as follows:

$$\hat{x}(t + \Delta t) = x(t) + f(x(t))\Delta t + g(x(t))\sqrt{\Delta t}\eta(t) \quad (\text{A6})$$

$$x(t + \Delta t) = x(t) + \frac{1}{2} (f(x(t)) + f(\hat{x}(t + \Delta t))\Delta t + \frac{1}{2} (g(x(t)) + g(\hat{x}(t + \Delta t))\sqrt{\Delta t}\eta(t). \quad (\text{A7})$$

ORCID iDs

Forough Hassanibesheli  <https://orcid.org/0000-0001-6919-4358>

References

- [1] Haken H 1980 Synergetics *Naturwissenschaften* **67** 121–8
- [2] Meyers R A 2010 *Complex Systems in Finance and Econometrics* (Berlin: Springer)
- [3] Fieguth P 2017 An introduction to complex systems *Complex Systems and Archaeology* vol 10 (Berlin: Springer) pp 978–83
- [4] Holovatch Y, Kenna R and Thurner S 2017 Complex systems: physics beyond physics *Eur. J. Phys.* **38** 023002
- [5] Della Rossa F and DeLellis P 2020 Stochastic master stability function for noisy complex networks *Phys. Rev. E* **101** 052211
- [6] Rakshit S, Bera B K, Bollt E M and Ghosh D 2020 Intralayer synchronization in evolving multiplex hypernetworks: analytical approach *SIAM J. Appl. Dyn. Syst.* **19** 918–63
- [7] Thiel M, Kurths J, Romano M C, Károlyi G and Moura A 2010 *Nonlinear Dynamics and Chaos: Advances and Perspectives* (Berlin: Springer)
- [8] Friedrich R and Peinke J 1997 Description of a turbulent cascade by a Fokker-Planck equation *Phys. Rev. Lett.* **78** 863
- [9] Jannesar M, Sadeghi A, Meyer E and Jafari G 2019 A langevin equation that governs the irregular stick-slip nano-scale friction *Sci. Rep.* **9** 1–7
- [10] Hegger R and Stock G 2009 Multidimensional langevin modeling of biomolecular dynamics *J. Chem. Phys.* **130** 034106
- [11] Ghasemi F, Sahimi M, Peinke J, Friedrich R, Jafari G R and Tabar M R R 2007 Markov analysis and Kramers-Moyal expansion of nonstationary stochastic processes with application to the fluctuations in the oil price *Phys. Rev. E* **75** 060102
- [12] Kubo R 1966 The fluctuation-dissipation theorem *Rep. Prog. Phys.* **29** 255
- [13] Cui B and Zaccane A 2018 Generalized langevin equation and fluctuation-dissipation theorem for particle-bath systems in external oscillating fields *Phys. Rev. E* **97** 060102
- [14] Uhlenbeck G E and Ornstein L S 1930 On the theory of the Brownian motion *Phys. Rev.* **36** 823

- [15] Mori H 1965 Transport, collective motion, and Brownian motion *Prog. Theor. Phys.* **33** 423–55
- [16] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (New York: Oxford University Press)
- [17] Darve E, Solomon J and Kia A 2009 Computing generalized langevin equations and generalized Fokker–Planck equations *Proc. Natl Acad. Sci.* **106** 10884–9
- [18] Zwanzig R 1961 Memory effects in irreversible thermodynamics *Phys. Rev.* **124** 983
- [19] Kravtsov S, Kondrashov D and Ghil M 2005 Multilevel regression modeling of nonlinear processes: derivation and applications to climatic variability *J. Clim.* **18** 4404–24
- [20] Kondrashov D, Chekroun M D and Ghil M 2015 Data-driven non-Markovian closure models *Physica D* **297** 33–55
- [21] Penland C 1989 Random forcing and forecasting using principal oscillation pattern analysis *Mon. Weather Rev.* **117** 2165–85
- [22] Chen C, Cane M A, Henderson N, Lee D E, Chapman D, Kondrashov D and Chekroun M D 2016 Diversity, nonlinearity, seasonality, and memory effect in ENSO simulation and prediction using empirical model reduction *J. Clim.* **29** 1809–30
- [23] Friedrich R and Peinke J 1997 Statistical properties of a turbulent cascade *Physica D* **102** 147–55
- [24] Friedrich R, Peinke J, Sahimi M and Tabar M R R 2011 Approaching complexity by stochastic methods: from biological systems to turbulence *Phys. Rep.* **506** 87–162
- [25] Nadaraya E A 1964 On estimating regression *Theor. Probab. Appl.* **9** 141–2
- [26] Watson G S 1964 Smooth regression analysis *Sankhya: Indian J. Stat., Series A* **26** 359–72
- [27] Lade S J 2009 Finite sampling interval effects in Kramers–Moyal analysis *Phys. Lett. A* **373** 3705–9
- [28] Lehnertz K, Zabawa L and Tabar M R R 2018 Characterizing abrupt transitions in stochastic dynamics *New J. Phys.* **20** 113043
- [29] Honisch C and Friedrich R 2011 Estimation of Kramers–Moyal coefficients at low sampling rates *Phys. Rev. E* **83** 066701
- [30] Friedrich R, Renner C, Siefert M and Peinke J 2002 Comment on indispensable finite time corrections for Fokker–Planck equations from time series data *Phys. Rev. Lett.* **89** 149401
- [31] Hijón C, Español P, Vanden-Eijnden E and Delgado-Buscalioni R 2010 Mori–Zwanzig formalism as a practical computational tool *Faraday Discuss.* **144** 301–22
- [32] Schmitt D T and Schulz M 2006 Analyzing memory effects of complex systems from time series *Phys. Rev. E* **73** 056204
- [33] Niemann M, Laubrich T, Olbrich E and Kantz H 2008 Usage of the Mori–Zwanzig method in time series analysis *Phys. Rev. E* **77** 011117
- [34] Li Z, Lee H S, Darve E and Karniadakis G E 2017 Computing the non-Markovian coarse-grained interactions derived from the Mori–Zwanzig formalism in molecular systems: application to polymer melts *J. Chem. Phys.* **146** 014104
- [35] Feldstein A and Sopka J R 1974 Numerical methods for nonlinear volterra integro-differential equations *SIAM J. Numer. Anal.* **11** 826–46
- [36] Vanualailai J and Nakagiri S-i 2003 Stability of a system of volterra integro-differential equations *J. Math. Anal. Appl.* **281** 602–19
- [37] Alexander M A, Matrosova L, Penland C, Scott J D and Chang P 2008 Forecasting Pacific SSTs: linear inverse model predictions of the PDO *J. Clim.* **21** 385–402
- [38] Horsthemke W 1984 Noise induced transitions *Non-Equilibrium Dynamics in Chemical Systems* (Berlin: Springer) pp 150–60
- [39] Sura P, Newman M, Penland C and Sardeshmukh P 2005 Multiplicative noise and non-gaussianity: a paradigm for atmospheric regimes? *J. Atmos. Sci.* **62** 1391–409
- [40] Perelló J, Porra J, Montero M and Masoliver J 2000 Black–scholes option pricing within Itô and Stratonovich conventions *Physica A* **278** 260–74
- [41] Gopikrishnan P, Plerou V, Liu Y, Amaral L N, Gabaix X and Stanley H E 2000 Scaling and correlation in financial time series *Physica A* **287** 362–73
- [42] Lahmiri S 2014 Multi-scaling analysis of the sp500 under different regimes in wavelet domain *Int. J. Strateg. Decis. Sci.* **5** 43–55
- [43] Rayner N, Parker D E, Horton E, Folland C K, Alexander L V, Rowell D, Kent E and Kaplan A 2003 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res. Atmos.* **108** 4407
- [44] Johnson S D, Battisti D S and Sarachik E 2000 Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies *J. Clim.* **13** 3–17
- [45] Alexander M A, Bladé I, Newman M, Lanzante J R, Lau N-C and Scott J D 2002 The atmospheric bridge: the influence of ENSO teleconnections on air–sea interaction over the global oceans *J. Clim.* **15** 2205–31
- [46] Kondrashov D, Kravtsov S, Robertson A W and Ghil M 2005 A hierarchy of data-based ENSO models *J. Clim.* **18** 4425–44
- [47] Chekroun M D, Kondrashov D and Ghil M 2011 Predicting stochastic systems by noise sampling, and application to the El Niño–southern oscillation *Proc. Natl Acad. Sci.* **108** 11766–71
- [48] Martínez-Villalobos C, Newman M, Vimont D J, Penland C and David Neelin J 2019 Observed El Niño–La Niña asymmetry in a linear model *Geophys. Res. Lett.* **46** 9909–19
- [49] Kwasniok F 2013 Analysis and modelling of glacial climate transitions using simple dynamical systems *Phil. Trans. R. Soc. A* **371** 20110472
- [50] Boers N, Chekroun M D, Liu H, Kondrashov D, Rousseau D-D, Svensson A, Bigler M and Ghil M 2017 Inverse stochastic–dynamic models for high-resolution greenland ice core records *Earth Syst. Dyn.* **8** 1171–90
- [51] Lohmann J and Ditlevsen P D 2019 A consistent statistical model selection for abrupt glacial climate changes *Clim. Dyn.* **52** 6411–26
- [52] Boers N, Ghil M and Rousseau D-D 2018 Ocean circulation, ice shelf, and sea ice interactions explain Dansgaard–Oeschger cycles *Proc. Natl Acad. Sci.* **115** E11005–14
- [53] Boers N 2018 Early-warning signals for Dansgaard–Oeschger events in a high-resolution ice core record *Nat. Commun.* **9** 1–8
- [54] Kinder P, Guillevic M, Baumgartner M F, Schwander J, Landais A and Leuenberger M 2014 Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core *Clim. Past* **10** 887–902
- [55] Rasmussen S O et al 2014 A stratigraphic framework for abrupt climatic changes during the last glacial period based on three synchronized greenland ice-core records: refining and extending the intimate event stratigraphy *Quat. Sci. Rev.* **106** 14–28
- [56] AghaKouchak A, Easterling D, Hsu K, Schubert S and Sorooshian S 2012 *Extremes in a Changing Climate: Detection, Analysis and Uncertainty* vol 65 (Berlin: Springer)
- [57] Garcia-Alvarez D 2011 A comparison of a few numerical schemes for the integration of stochastic differential equations in the Stratonovich interpretation (arXiv:1102.4401)
- [58] Stratonovich R L 1967 *Topics in the Theory of Random Noise* vol 2 (Boca Raton, FL: CRC Press)
- [59] Kloeden P E and Platen E 1991 Stratonovich and Itô Stochastic Taylor expansions *Math. Nachr.* **151** 33–50
- [60] Shreve S 2005 *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model* (Berlin: Springer)
- [61] Moon W and Wettlaufer J 2014 On the interpretation of Stratonovich calculus *New J. Phys.* **16** 055017

- [62] Kupferman R, Pavliotis G A and Stuart A M 2004 Itô versus stratonovich white-noise limits for systems with inertia and colored multiplicative noise *Phys. Rev. E* **70** 036120
- [63] Jacobs K 2010 *Stochastic Processes for Physicists: Understanding Noisy Systems* (Cambridge: Cambridge University Press)
- [64] Rümelin W 1982 Numerical treatment of stochastic differential equations *SIAM J. Numer. Anal.* **19** 604–13
- [65] Greiner A, Strittmatter W and Honerkamp J 1988 Numerical integration of stochastic differential equations *J. Stat. Phys.* **51** 95–108